

On the strong law of large numbers for φ -subgaussian random variables

KRZYSZTOF ZAJKOWSKI

Institute of Mathematics, University of Białystok
Ciołkowskiego 1M, 15-245 Białystok, Poland
kryza@math.uwb.edu.pl

Abstract

For $p \geq 1$ let $\varphi_p(x) = x^2/2$ if $|x| \leq 1$ and $\varphi_p(x) = 1/p|x|^p - 1/p + 1/2$ if $|x| > 1$. For a random variable ξ let $\tau_{\varphi_p}(\xi)$ denote $\inf\{a \geq 0 : \forall \lambda \in \mathbb{R} \ln \mathbb{E} \exp(\lambda \xi) \leq \varphi_p(a\lambda)\}$; τ_{φ_p} is a norm in a space $Sub_{\varphi_p} = \{\xi : \tau_{\varphi_p}(\xi) < \infty\}$ of φ_p -subgaussian random variables. We prove that if for a sequence $(\xi_n) \subset Sub_{\varphi_p}$ ($p > 1$) there exist positive constants c and α such that for every natural number n the following inequality $\tau_{\varphi_p}(\sum_{i=1}^n \xi_i) \leq cn^{1-\alpha}$ holds then $n^{-1} \sum_{i=1}^n \xi_i$ converges almost surely to zero as $n \rightarrow \infty$. This result is a generalization of the SLLN for independent subgaussian random variables (Taylor and Hu [9]) to the case of dependent φ_p -subgaussian random variables.

2010 Mathematics Subject Classification: 60F15

Key words: φ -subgaussian random variables, strong law of large numbers

1 Introduction

The classical Kolmogorov's strong laws of large numbers are dealt with independent variables. Investigations of limit theorems for dependent r.v.s are extensive and episodic. The strong law of large numbers for various classes of many type associated random variables one can find for instance in Bulinski and Shashkin [2, Chap. 4]. Most of them are considered in the spaces of integrable functions. It is also interested to describe general conditions under which the SLLN holds in other spaces of random variables than L_p -spaces. In this paper we investigate almost sure convergence of the arithmetic mean (but not only) sequences of φ -subgaussian random variables.

The notion of subgaussian random variables was introduced by Kahane in [8]. A random variable ξ is *subgaussian* (σ -*subgaussian*) if its moment generating function is majorized by the moment generating function of some centered gaussian r.v. with variance σ^2 that is $\mathbb{E} \exp(\lambda \xi) \leq \mathbb{E} \exp(\lambda g) = \exp(\sigma^2 \lambda^2 / 2)$, where $g \sim \mathcal{N}(0, \sigma^2)$ (see Buldygin and Kozachenko [4] or [3, Ch.1]). In terms of the cumulant generating functions this condition takes a form: $\ln \mathbb{E} \exp(\lambda \xi) \leq \sigma^2 \lambda^2 / 2$.

One can generalize the notion of subgaussian r.v.s to classes of φ -subgaussian random variables (see [3, Ch.2]). A continuous even convex function $\varphi(x)$ ($x \in \mathbb{R}$) is called a *N-function*, if the following condition hold:

- (a) $\varphi(0) = 0$ and $\varphi(x)$ is monotone increasing for $x > 0$,
(b) $\lim_{x \rightarrow 0} \varphi(x)/x = 0$ and $\lim_{x \rightarrow \infty} \varphi(x)/x = \infty$.

It is called a *quadratic N -function*, if in addition $\varphi(x) = cx^2$ for all $|x| \leq x_0$, with $c > 0$ and $x_0 > 0$. The quadratic condition is needed to ensure nontriviality for classes of φ -subgaussian random variables (see [3, Ch.2, p.67]).

Example 1.1. Let for $p \geq 1$

$$\varphi_p(x) = \begin{cases} \frac{x^2}{2}, & \text{if } |x| \leq 1, \\ \frac{1}{p}|x|^p - \frac{1}{p} + \frac{1}{2}, & \text{if } |x| > 1. \end{cases}$$

The function φ_p is an example of the quadratic N -function which is some standardization of the function $(1/p)|x|^p$. Let us emphasize that for $p = 2$ we have the case of subgaussian random variables.

Let φ be a quadratic N -function. A random variable ξ is said to be φ -subgaussian if there is a constant $a > 0$ such that $\ln \mathbb{E} \exp(\lambda \xi) \leq \varphi(a\lambda)$. The φ -subgaussian standard (norm) $\tau_\varphi(\xi)$ is defined as

$$\tau_\varphi(\xi) = \inf\{a \geq 0 : \forall \lambda \in \mathbb{R} \quad \ln \mathbb{E} \exp(\lambda \xi) \leq \varphi(a\lambda)\};$$

a space $Sub_\varphi = \{\xi : \tau_\varphi(\xi) < \infty\}$ with the norm τ_φ is a Banach space (see [3, Ch.2, Th.4.1])

Let $\varphi(x)$ ($x \in \mathbb{R}$) be a real-valued function. The function $\varphi^*(y)$ ($y \in \mathbb{R}$) defined by $\varphi^*(y) = \sup_{x \in \mathbb{R}} \{xy - \varphi(x)\}$ is called the *Young-Fenchel transform* or the *convex conjugate* of φ (in general, φ^* may take value ∞). It is known that if φ is a quadratic N -function then φ^* is quadratic N -function too. For instance, since our φ_p is a differentiable (even at ± 1) function one can easily check that $\varphi_p^* = \varphi_q$ for $p, q > 1$, if $1/p + 1/q = 1$.

Example 1.2. Let $g \sim \mathcal{N}(0, 1)$ then $\xi = |g|^{2/q} - \mathbb{E}|g|^{2/q} \in Sub_{\varphi_p}(\Omega)$, where $1/p + 1/q = 1$.

Let us recall that the convex conjugate is order-reversing and possesses some scaling property. If $\varphi_1 \geq \varphi_2$ then $\varphi_1^* \leq \varphi_2^*$. Let for $a > 0$ and $b \neq 0$ $\psi(x) = a\varphi(bx)$ then $\psi^*(y) = a\varphi^*(y/(ab))$ (see e.g. [6, Ch.X, Prop.1.3.1]).

The convex conjugate of the cumulant generating function can be served to estimate of 'tails' distribution of a centered random variable. Let $\mathbb{E}\xi = 0$ and ψ_ξ denote the cumulant generating function of ξ , i.e. $\psi_\xi(\lambda) = \ln \mathbb{E} \exp(\lambda \xi)$ then for $\varepsilon > 0$

$$\mathbb{P}(\xi \geq \varepsilon) \leq \exp(-\psi_\xi^*(\varepsilon)).$$

Let us observe that for $\xi \in Sub_\varphi$, by the definition of $\tau_\varphi(\xi)$, we have the following inequality: $\psi_\xi(\lambda) \leq \varphi(\tau_\varphi(\xi)\lambda)$ and by the order-reversing and the scaling property we

get $\psi_\xi^*(\varepsilon) \geq \varphi^*(\varepsilon/\tau_\varphi(\xi))$. Now we can obtain some weaker form of the above estimation but with using the general function φ :

$$\mathbb{P}(|\xi| \geq \varepsilon) \leq 2 \exp \left(-\varphi^* \left(\frac{\varepsilon}{\tau_\varphi(\xi)} \right) \right); \quad (1)$$

see [3, Ch.2, Lem.4.3].

2 Results

First we show that if we have some upper estimate for τ_φ then in (1) we can substitute this estimate instead of τ_φ .

Lemma 2.1. *If $\tau_\varphi(\xi) \leq C(\xi)$ for every $\xi \in \text{Sub}_\varphi$ then*

$$\mathbb{P}(|\xi| \geq \varepsilon) \leq 2 \exp \left(-\varphi^* \left(\frac{\varepsilon}{C(\xi)} \right) \right).$$

Proof. Since φ is even and increasing monotonic for $x > 0$, we get

$$\varphi(\tau_\varphi(\xi)x) = \varphi(\tau_\varphi(\xi)|x|) \leq \varphi(C(\xi)|x|) = \varphi(C(\xi)x).$$

And again by the order-reversing and the scaling property we obtain

$$\varphi^* \left(\frac{y}{\tau_\varphi(\xi)} \right) \geq \varphi^* \left(\frac{y}{C(\xi)} \right),$$

which combined with (1) establishes the inequality. \square

With these preliminaries accounted for, we can prove the main result of the paper.

Theorem 2.2. *Let $(\xi_n) \subset \text{Sub}_{\varphi_p}$ for some $p > 1$. If there exist positive constants c and α such that for every natural number n the following condition $\tau_{\varphi_p}(\sum_{i=1}^n \xi_i) \leq cn^{1-\alpha}$ holds then the term $n^{-1} \sum_{i=1}^n \xi_i$ converges almost surely to zero as $n \rightarrow \infty$.*

Proof. Since $\varphi_p^* = \varphi_q$, by Lemma 2.1 and the condition of the theorem we have

$$\mathbb{P} \left(\left| \sum_{i=1}^n \xi_i \right| \geq n\varepsilon \right) \leq 2 \exp \left(-\varphi_q \left(\frac{n^\alpha \varepsilon}{c} \right) \right).$$

For sufficiently large n ($n > (c/\varepsilon)^{1/\alpha}$) we have $n^\alpha \varepsilon / c > 1$ and, in consequence,

$$\varphi_q \left(\frac{n^\alpha \varepsilon}{c} \right) = n^{q\alpha} \frac{1}{q} \left(\frac{\varepsilon}{c} \right)^q - \frac{1}{q} + \frac{1}{2}.$$

Thus we get the following estimate

$$\mathbb{P}\left(\left|\sum_{i=1}^n \xi_i\right| \geq n\varepsilon\right) \leq 2 \exp\left(\frac{1}{q} - \frac{1}{2}\right) \exp\left(-n^{q\alpha} \frac{1}{q} \left(\frac{\varepsilon}{c}\right)^q\right)$$

for every ε and $n > (c/\varepsilon)^{1/\alpha}$. Thus, by the integral test, we obtain convergence of the series $\sum_{n=1}^{\infty} \mathbb{P}(|\sum_{i=1}^n \xi_i| \geq n\varepsilon)$. It follows the completely and, in consequence, almost sure convergence of $n^{-1} \sum_{i=1}^n \xi_i$ to zero. \square

Remark 2.3. Let us emphasize that the above theorem is a generalization of the Theorem (SLLN) (Taylor and Hu [9, sec.3, p.297]) to the case of φ_p -subgaussian random variables, not only subgaussian ones. Moreover we do not assume their independence. For this reason we used a modified condition for a behavior of the norm τ_p than Taylor and Hu, which I describe below.

Since τ_φ is a norm, we obtain

$$\tau_\varphi\left(\sum_{i=1}^n \xi_i\right) \leq \sum_{i=1}^n \tau_\varphi(\xi_i).$$

If for instance ξ_i , $i = 1, \dots, n$, are copies of the same variable ξ then in the above the equality holds and $\tau_\varphi\left(\sum_{i=1}^n \xi_i\right) = n\tau_\varphi(\xi)$. Let us observe that in this case the assumption of Theorem 2.2 is not satisfied. Additionally informations about form of dependence (or independence) sometime allow us to improve this estimate. And so, for an independence sequence (ξ_n) if there is some $r \in (0, 2]$ such that $\varphi(|x|^{1/r})$ is convex then

$$\tau_\varphi\left(\sum_{i=1}^n \xi_i\right)^r \leq \sum_{i=1}^n \tau_\varphi(\xi_i)^r; \quad (2)$$

see [3, Sec.2, Th.5.2]. If r is bigger then the estimate is better. For the function φ_p we can always take $r = \min\{p, 2\}$. In Taylor's and Hu's SLLN variables ξ_n were subgaussian and independent and it was taken $p = 2$. Let us emphasize that in this case if in addition ξ_1, \dots, ξ_n have the same distribution as ξ then $\tau_\varphi(\sum_{i=1}^n \xi_i) \leq \sqrt{n}\tau_\varphi(\xi)$ and the condition of Theorem 2.2 is satisfied ($c = \tau_\varphi(\xi)$ and $\alpha = 1/2$).

Let us emphasize that another assumptions on dependence of ξ_1, \dots, ξ_n can give the same estimate of the norm of $\tau_\varphi(\sum_{i=1}^n \xi_i)$. In the paper Giuliano Antonini et al.[5, Lem.3] it was proved that for φ -subgaussian acceptable random variables the inequality (2) holds, if $\varphi(|x|^{1/r})$ is convex. The definition of acceptability of sequence of random variable one can find therein. For us it is the most important that these estimates are the same. In this article there is some version of the Marcinkiewicz-Zygmund law of large numbers for φ -subgaussian random variables as a corollary of much more

general theorem. We give an independent proof of this corollary but under modified assumptions.

Proposition 2.4. *For $p > 1$ let (ξ_n) be a bounded sequence of φ_p -subgaussian random variables and let $r = \min\{p, 2\}$. If in addition*

$$\tau_{\varphi_p} \left(\sum_{i=1}^n \xi_i \right)^r \leq \sum_{i=1}^n \tau_{\varphi_p}(\xi_i)^r \quad (3)$$

then $n^{-1/s} \sum_{i=1}^n \xi_i \rightarrow 0$ almost surely for any $0 < s < r$.

Remark 2.5. Since $\varphi_p(|x|^{1/r})$ is convex, the estimate (3) is satisfied by sequences of independent or acceptable random variables, for instance.

Proof. Let $b = \sup_{n \geq 1} \tau_{\varphi_p}(\xi_n)$ then $\sum_{i=1}^n \tau_{\varphi_p}(\xi_i)^r \leq nb^r$ and, in consequence, $\tau_{\varphi_p} \left(\sum_{i=1}^n \xi_i \right) \leq n^{1/r}b$. For positive number s less than r , by Lemma 2.1, we obtain

$$\mathbb{P} \left(\left| \sum_{i=1}^n \xi_i \right| \geq n^{\frac{1}{s}} \varepsilon \right) \leq 2 \exp \left(- \varphi_q \left(\frac{n^{1/s} \varepsilon}{n^{1/r} b} \right) \right) = 2 \exp \left(- \varphi_q \left(n^{(\frac{1}{s} - \frac{1}{r})} \frac{\varepsilon}{b} \right) \right).$$

For $n > (b/\varepsilon)^{(1/s - 1/r)^{-1}}$ we have

$$\varphi_q \left(n^{(\frac{1}{s} - \frac{1}{r})} \frac{\varepsilon}{b} \right) = n^{q(\frac{1}{s} - \frac{1}{r})} \frac{1}{q} \left(\frac{\varepsilon}{b} \right)^q - \frac{1}{q} + \frac{1}{2}$$

and, in consequence, we get

$$\sum_{n=1}^{\infty} \exp \left(- \varphi_q \left(n^{(\frac{1}{s} - \frac{1}{r})} \frac{\varepsilon}{b} \right) \right) < \infty,$$

which, in view of Borel-Cantelli lemma, completes the proof. \square

Remark 2.6. Because we apply the function $\varphi_p(x)$ instead of $1/p|x|^p$ then we must not restrict p to be less or equal 2 to ensure the fulfillment of the quadratic condition for the function $1/p|x|^p$. Moreover we use the metric property (3) instead of assumptions on some form of dependence random variables (compare [5, Cor. 7]).

Example 2.7. The proof of Hoeffding-Azuma's inequality for a sequence (ξ_n) of bounded random variables such that $|\xi_n| \leq d_n$ a.s. and $\mathbb{E}\xi_n = 0$ is based on an estimate of the moment generating function of the partial sum $\sum_{i=1}^n \xi_i$. Under assumptions that ξ_n are independent (Hoeffding) or ξ_n are martingales increments (Azuma) the following inequality holds

$$\mathbb{E} \exp \left(\lambda \sum_{i=1}^n \xi_i \right) \leq \exp \left(\frac{\lambda^2 \sum_{i=1}^n d_i^2}{2} \right); \quad (4)$$

see Hoeffding [7] and Azuma [1]. Let us emphasize that in [1] Azuma has proved the above estimate under more general assumptions on (ξ_n) which satisfy centered bounded martingales increments. The inequality (4) means that

$$\tau_{\varphi_2}\left(\sum_{i=1}^n \xi_i\right) \leq \left(\sum_{i=1}^n d_i^2\right)^{1/2}.$$

If we take $d_n = 1$ for $n = 1, 2, \dots$ then we get the following condition

$$\tau_{\varphi_2}\left(\sum_{i=1}^n \xi_i\right) \leq \sqrt{n},$$

which follows that the sequence (ξ_n) satisfies the assumptions of Proposition 2.4 with $p = r = 2$ and the norm $\tau_{\varphi_2}(\xi_n) \leq 1$ and we get the almost sure convergence $n^{-1/s} \sum_{i=1}^n \xi_i$ to 0 for any $0 < s < 2$. Let us note that for $s = 1$ we obtain SLLN for this sequence.

References

- [1] K. Azuma, *Weighted sums of certain dependent random variables*, Tokohu Mathematical Journal 19 (1967), 357-367.
- [2] A. Bulinski, A. Shashkin, *Limit Theorems for Associated Random Fields and Related Systems*, World Scientific Publishing Co. Pte. Ltd. 2007.
- [3] V. Buldygin, Yu. Kozachenko, *Metric Characterization of Random Variables and Random Processes*, Amer.Math.Soc., Providence, RI, 2000.
- [4] V. Buldygin, Yu. Kozachenko, *Subgaussian random variables*, Ukrainian Math. J. 32 (1980), 483-489.
- [5] R. Giuliano Antonini, Yu. Kozaczenko, A. Volodin, *Convergence of series of dependent φ -subgaussian random variables*, J. Math. Anal. Appl. 338(2008) 1188-1203.
- [6] J.-B. Hiriart-Urruty, C. Lemaréchal, *Convex Analysis and Minimization Algorithms. II*, Springer-Verlag, Berlin Heidelberg 1993.
- [7] W. Hoeffding, *Probability for sums of bounded random variables*, Journal of the American Statistical Association 58 (1963), 13-30.
- [8] J.P. Kahane, *Local properties of functions in terms of random Fourier series (French)*, Stud. Math., 19 (no. 1), 1-25 (1960).
- [9] R.L. Taylor, T.-C. Hu, *Sub-Gaussian techniques in proving strong laws of large numbers*, Amer. Math. Monthly 94 (1987) 295-299.